WO0207438

Publication Title:

GENERALIZED LAPPED BIORTHOGONAL TRANSFORM EMBEDDED INVERSE DISCRETE COSINE TRANSFORM AND LOW BIT RATE VIDEO SEQUENCE CODING ARTIFACT REMOVAL

Abstract:

Abstract of WO0207438

A generalized lapped biorthogonal transform embedded inverse discrete cosine transform (ge-IDCT) (20), as an alternative to the inverse discrete cosine transform (IDCT) within a system for still image compression. The ge-IDCT (20) takes advantage of the DCT (18) front end of the generalized lapped biorthogonal transform (GLBT) in inverse transforming the DCT (18) coefficients. Non-linear weighting is used in the embedded lapped transform domain (16), so that the ge-IDCT (20) can reconstruct the signal with alleviated blockishness. In another embodiment, the disclosed system includes a post-processing method to reduce anomalies caused by blocking artifacts by applying a lapped orthogonal transform-embedded inverse discrete cosine transform (le-IDCT), as a substitute for the usual inverse DCT (18). For the reduction of ringing artifacts, a nonlinear robust filter is applied to the decoded picture frame. Data supplied from the esp@cenet database - Worldwide

----

Courtesy of http://v3.espacenet.com

(51) International Patent Classification[7]: **H04N 7/12**

(21) International Application Number: PCT/US01/22368

(22) International Filing Date: 17 July 2001 (17.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/218,600   17 July 2000 (17.07.2000)   US

(71) Applicant *(for all designated States except US)*: TRUSTEES OF BOSTON UNIVERSITY [US/US]; 147 Bay State Road, Boston, MA 02215 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: NGUYEN, Troung, Quang [US/US]; Apartment 5, 7718 Magnolia Avenue, Riverside, CA 92504 (US). YANG, Seungjoon [KR/KR]; Byuk Jeok Gol, Joo Gong Apt. 834-2001, Youngtong-Dong, Paldal-ku, Suwon City, Kyungki-Do 442-470 (KR).

KITTITORNKUN, Surin [TH/US]; Apartment 208, 2020 University Avenue, Madison, WI 53705 (US). HU, Yu, Hen [US/US]; 1640 Foxridge Court, Middleton, WI 53562 (US). TULL, Damon, Lee [US/US]; 20 Saynor Court, #44, Madison, WI 53717 (US).

(74) Agents: GAGNEBIN, Charles, L., III et al.; Weingarten, Schurgin, Gagnebin & Hayes, LLP, Ten Post Office Square, Boston, MA 02109 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,

(54) Title: GENERALIZED LAPPED BIORTHOGONAL TRANSFORM EMBEDDED INVERSE DISCRETE COSINE TRANSFORM AND LOW BIT RATE VIDEO SEQUENCE CODING ARTIFACT REMOVAL
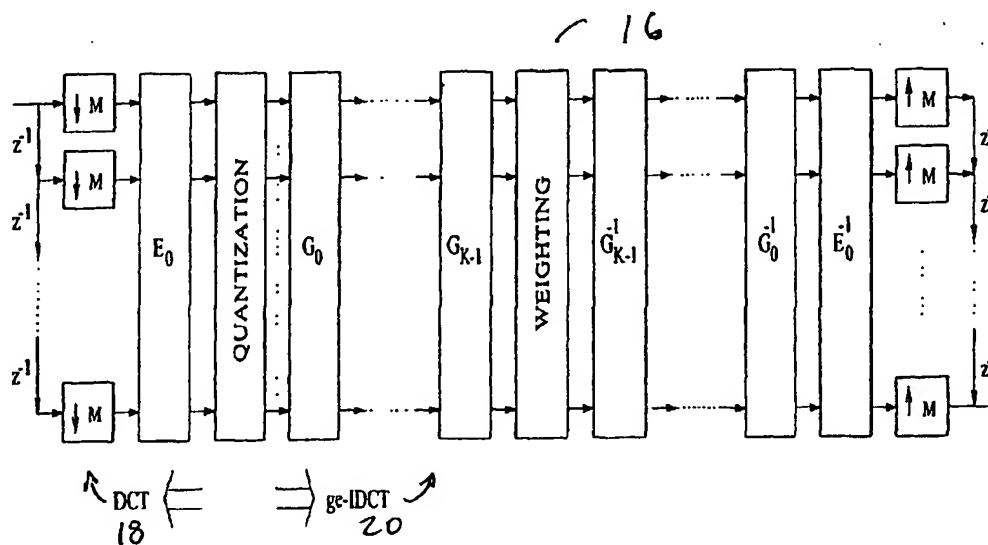
(57) Abstract: A generalized lapped biorthogonal transform embedded inverse discrete cosine transform (ge-IDCT) (20), as an alternative to the inverse discrete cosine transform (IDCT) within a system for still image compression. The ge-IDCT (20) takes advantage of the DCT (18) front end of the generalized lapped biorthogonal transform (GLBT) in inverse transforming the DCT (18) coefficients. Non-linear weighting is used in the embedded lapped transform domain (16), so that the ge-IDCT (20) can reconstruct the signal with alleviated blockishness. In another embodiment, the disclosed system includes a post-processing method to reduce anomalies caused by blocking artifacts by applying a lapped orthogonal transform-embedded inverse discrete cosine transform (le-IDCT), as a substitute for the usual inverse DCT (18). For the reduction of ringing artifacts, a nonlinear robust filter is applied to the decoded picture frame.

CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*
— *of inventorship (Rule 4.17(iv)) for US only*
— *of inventorship (Rule 4.17(iv)) for US only*

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

TITLE
Generalized Lapped Biorthogonal Transform
Embedded Inverse Discrete Cosine Transform and Low Bit
Rate Video Sequence Coding
Artifact Removal

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims priority under 35 U.S.C. §119(e) to provisional patent application serial number 60/218,600, entitled IMAGE BLOCKING ARTIFACT REDUCTION USING LAPPED ORTHOGONAL TRANSFORM EMBEDDED INVERSE DISCREET COSINE TRANSFORM, filed July 17, 2000.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT
N/A

BACKGROUND OF THE INVENTION

Most of the existing still image and video sequence compression standards employ a block based discrete cosine transform (DCT). Further in existing still image and video sequence compression systems, at mid to low bit rates, picture quality deteriorates due to the presence of coding artifacts. The use of the block based DCT often leads to an annoying coding artifact known as the *blocking* artifact, which exhibits itself as visible discontinuity at block boundaries. The blocking artifact

is due to the short basis functions of the DCT and
independent processing of blocks. Lapped transforms, in
which the data blocks overlap each other, have been
introduced to reduce or eliminate the blocking artifact.

5   Another type of coding artifact, known as the
ringing artifact, exhibits itself as spurious
oscillations around the vicinity of major edges of the
image. The ringing artifact is due to abrupt truncation
of high frequency components. The ringing artifact is
10   also known as mosquito noise.

Picture quality can be enhanced by various methods
of post-processing. Existing post-processing approaches
include MAP estimation, projection onto convex sets
(POCS), and linear/nonlinear filtering. MAP estimation
15   and POCS based algorithms are iterative and complicated
algorithms. Each step involves forward and inverse
transforms due to constraints in different domains. The
high computational complexity of these algorithms
prohibits their application to real time video sequence
20   decoding. Existing filtering based post-processing
algorithms involve a number of decision steps to detect
the occurrence, level, and type of degradation, and to
choose the corresponding filter for enhancement.
Propagation of these decision steps to the next frame is
25   often required.

Accordingly, it would be desirable to have a new
method for reducing the anomalies caused by blocking
and/or ringing artifacts, which avoids the problems
exhibited by previous systems, and which may be applied

-3-

to improve the picture quality of still images and/or video frames encoded at mid to low bit rates.

BRIEF SUMMARY OF THE INVENTION

5

The disclosed system includes a generalized lapped biorthogonal transform embedded inverse discrete cosine transform (ge-IDCT), as an alternative to the inverse discrete cosine transform (IDCT) within a system for
10 still image compression. The ge-IDCT takes advantage of the DCT front end of the generalized lapped biorthogonal transform (GLBT), such that it can be used in inverse transforming the DCT coefficients. With the nonlinear weighting in the embedded lapped transform domain, the
15 ge-IDCT can reconstruct the signal with alleviated blockishness. Additional complexity imposed by the replacement of the IDCT by the ge-IDCT is trivial thanks to an efficient lattice structure. In an illustrative embodiment, the disclosed ge-IDCT may be applied in the
20 JPEG still image compression standard.

In another embodiment, the disclosed system improves the picture quality of video frames encoded at relatively low-bit rates by reducing the effects of both blocking and ringing artifacts. In this embodiment, the disclosed
25 system includes two picture post-processing methods to reduce the anomalies caused by these artifacts. For the reduction of blocking artifacts, the disclosed system operates to apply a lapped orthogonal transform-embedded inverse discrete cosine transform (le-IDCT), as a
30 substitute for the usual inverse DCT. In this way, the

disclosed system allows data samples from adjacent blocks
to be processed simultaneously so that existing blocking
artifacts can be efficiently mitigated.　For the
reduction of ringing artifacts, the disclosed system may

5　　be embodied to include a nonlinear robust filter to be
applied to the decoded picture frame.

　　　The disclosed system advantageously provides marked
improvement in terms of both objective and subjective
image quality.　The computation overhead incurred due to

10　　the disclosed procedures is quite moderate, and real-time
implementations may be embodied in hardware, software,
firmware, or some combination thereof, executing on
common desktop computer systems.

15　　　　BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS
　　　The invention will be more fully understood by
reference to the following detailed description of the
invention in conjunction with the drawings, of which:

　　　Fig. 1 shows a flowgraph of GLBT, in which the

20　　analysis FB and the synthesis FB represent the forward
and the inverse transforms, respectively;

　　　Fig. 2 shows a flowgraph of the DCT and the ge-IDCT,
in which the ge-IDCT works in the case where the signal
is processed in the DCT domain, and frequency weighting

25　　is employed in the embedded GLBT domain;

　　　Fig. 3 shows the detailed lattice structure of an
analysis FB, including the first stage (a) with a DCT
front end, and also showing each stage (b);

Fig. 4 shows the detailed lattice structure of a
synthesis FB, including the last stage (a) with an IDCT
rear end, and each stage (b);

Fig. 5 shows non-overlapping transforms (a) and
overlapping transforms (b);

Fig. 6 illustrates improvement of the ge-IDCT with
nonlinear weighting in PSNR at various quality factors,
(PSNR of the proposed methods) - (PSNR of JPEG) in [dB]
vs. quality factor, for (a) airplane, (b) Barbara, (c)
Lena, and (d) peppers images;

Fig. 7 illustrates improvement of the ge-IDCT with
nonlinear weighting in MSDS at various quality factors,
(MSDS of the proposed methods) - (MSDS of JPEG) in [dB]
vs. quality factor, for (a) airplane, (b) Barbara, (c)
Lena, and (d) peppers images;

Fig. 8 shows a portion (a) of the Lena test image,
compressed by JPEG at quality factor 15, and an
associated edgemap (b);

Fig. 9 shows blocking artifact removal by MAP
estimation and the proposed method, including image (a)
by the MAP estimation, edge map (b) by the MAP estimation
with line process, image (c) by the proposed
method, and edge map (d) by the proposed method, and
further showing that most of the texture in Lena's hat is
missing in the MAP estimate due to over-smoothing;

Fig. 10 shows test image (a) and edge map (b) as
compressed by H.263I Frame coding method at QP = 13, and
reflecting by H.263 I Frame compression;

Fig. 11 illustrates blocking artifact removal by a
deblocking filter and by the disclosed method, where (a)

shows the image by the deblocking filter, (b) shows the edge map generated by the deblocking filter, (c) shows image generated by the disclosed method, and (d) shows edge map generated by the proposed method, and showing that the image processed by the deblocking filter is still relatively blockish due to under-smoothing;

Fig. 12 shows a design example of the GenLOT, including (a) impulse response, and (b) frequency response;

Fig. 13 shows detailed lattice structure of an embodiment of the disclosed le-IDCT;

Fig. 14 shows schematics of a modified video sequence decoder, wherein "le" represents a part of the le-IDCT that precedes the IDCT, "RF" represents the robust filter, and "s" is a switch;

Fig. 15 illustrates improvement obtained through the disclosed method in terms of PSNR, in dB, (PSNR of the proposed method) - (PSNR of baseline H.263+), at 24kb/s except *foreman* at 48 kb/s, I Frames every 100 frames, QP=13, for (a) *foreman*, and (b) *hall* images;

Fig. 16 illustrates improvement obtained through the disclosed method in terms of MSDS, in dB, (MSDS of the proposed method) - (MSDS of baseline H.263+), at 24kb/s except *foreman* at 48 kb/s, I Frames every 100 frames, QP=13, for (a) *foreman*, (b) *hall* images;

Fig. 17 shows video coding artifact removal by the proposed method, the 68th frame of *hall* sequence, compressed at 24kb/s, I Frame every 100 frames, QP=13, (a) image by baseline H.263+, (b) image by the disclosed method, (c) edge map by baseline H.263+, and (d) edge map

by the proposed method, showing effective removal of artifacts;

Fig. 18 shows video coding artifact removal by various algorithms, compressed at 24kb/s, I Frame every 100 frames, QP=13, with options in Annex J, (a) image by H.263+ deblocking filter, (b) image by the disclosed method, (c) edge map by H.263+ deblocking filter, and (d) edge map by the disclosed method, illustrating that the H.263+ deblocking filter can not remove the blockishness completely, and that the disclosed method provides effective removal of artifacts;

Fig. 19 is a table showing PSNR of compressed and processed sequences, in dB, at 24kb/s, I Frames every 100 frames, wherein results of the disclosed system are represented by the "proposed" column values;

Fig. 20 is a table showing MSDS of compressed and processed sequences, at 24kb/s, I Frames every 100 frames, wherein results of the disclosed system are represented by the "proposed" column values;

Fig. 21 is a table showing comparison of video coding artifact removal algorithms in PNSR, in dB, at 24kb/s, I Frames every 100 frames, (DF: deblocking filter in Annex J) , wherein results of the disclosed system are represented by the "proposed" column values;

Fig. 22 is a table showing comparison of video coding artifact removal algorithms in MSDS, at 24kb/s, I Frames every 100 frames, (DF: deblocking filter in Annex J), wherein results of the disclosed system are represented by the "proposed" column values;

Fig. 23 is a table showing comparison of average run time complexity of video coding artifact removal algorithms on I Frame, in [sec], (DF: deblocking filter in Annex J), wherein results of the disclosed system are

5   represented by the "proposed" column values; and
Fig. 24 is a table showing comparison of average run time complexity of video coding artifact removal algorithms on P Frame, in [sec], (DF: deblocking filter in Annex J), wherein results of the disclosed system are represented

10   by the "proposed" column values.

DETAILED DESCRIPTION OF THE INVENTION

United States Provisional Patent Application Serial

15   No. 60/218,600, entitled IMAGE BLOCKING ARTIFACT REDUCTION USING LAPPED ORTHOGONAL TRANSFORM EMBEDDED INVERSE DISCREET COSINE TRANSF, filed July 17, 2000, is hereby incorporated herein by reference.

20   I. Generalized Lapped Biorthogonal Transform Embedded Inverse Discrete Cosine Transform

The disclosed system embodies a method of utilizing a lapped transform in such a way that modification only

25   in the decoder section of existing systems is required. Existing encoders may be used without any modification to supply standard bit streams. The disclosed method is compliant with the current image/video compression standards that employ the DCT.

The generalized lapped biorthogonal transform (GLBT) is the most general form of lapped transforms. The GLBT is a linear phase perfect reconstruction filter bank (LPPRFB) based on the LP propagating lattice structure. The DCT is often used as the front end of the GLBT for its fast and efficient implementations. The DCT front end allows the GLBT to be used in inverse transforming the DCT coefficients. The DCT coefficients can be regarded as intermediate results of the GLBT with the DCT front end. Hence, the disclosed system may complete the rest of the stages in the analysis filter bank (FB), followed by the synthesis FB to reconstruct the signal. This operation is called the GLBT embedded inverse DCT (ge-IDCT).

The disclosed ge-IDCT provides an excellent opportunity to process the signal. In the case where the DCT coefficients are processed already, by the quantization operation for example, the signal can be re-processed in the embedded lapped transform domain to abate impairment of image quality. The blocking artifacts in image/video compression are degradation introduced by coarse quantizations of the DCT coefficients. In order to eliminate the blocking artifacts, the disclosed system employs nonlinear weighting of lapped transform coefficients.

The disclosed ge-IDCT with nonlinear weighting may be applied in the JPEG still image compression standard. The IDCT of the standard decoder is simply replaced by the ge-IDCT with nonlinear weighting. Experimental

results show consistent improvement of image quality at various bit rates.

Section I(A) below introduces the GLBT. Section I(B) below presents the disclosed ge-IDCT that can be paired with the forward DCT. Section I(C) presents the disclosed nonlinear weighting that reduces the blockishness in reconstructed images. Section V addresses the design of the ge-IDCT. In Section I(D), the ge-IDCT is applied to still image compression.

A. Generalized Lapped Biorthogonal Transform

The GLBT is a lapped transform defined as an LPPRFB with the polyphase transfer matrix (PTM), given by equation (1). The first stage E_0 is an M-channel LPPRFB with no delay element, which can be factored as shown in equation (2), in which I and J are the [M/2 x M/2] identity and reversal matrices, and the matrices U_0 and V_0 are [M/2 x N/2] invertible matrices. The PTM of each stage $G_i(z)$ is given by equation (3). The matrices U_i and V_i are [M/2 x N/2] invertible matrices. The matrix $\Lambda(z)$ has the delay element z^{-1}. The filter lengths increase by M by the delay element of each stage. The total length of the filter becomes KM. Note that the analysis FB is an M-channel FB, hence $(K - 1)M$ tabs of the filter lap over to the samples in previous blocks.

By using a singular value decomposition, the disclosed system factorizes each Ψ matrix as shown in

equation (4), where U_{ij} and V_{ij} are orthogonal matrices and Γ_i and Δ_i are diagonal matrices.

The PTM of the synthesis FB is given as shown in equation (5), such that the relationships in equation (6) hold true, and hence the perfect reconstruction (PR). The inverse matrices G_i^{-1} and E_0^{-1} involve the transposition of orthogonal matrices and inversion of diagonal matrices, which are trivial.

The matrices in the PTM are subject to design procedure. These matrices, or their equivalent Givens rotation angles, are optimized for better properties such as coding gain and stopband attenuation. The flowgraph 10 of the analysis FB 12 and the synthesis FB 14 of the GLBT are given in Fig. 1.

For $K > 1$, the data blocks of the GLBT overlap each other. Moreover, the basis functions of the GLBT have shapes that decay smoothly to zero. When the signal is processed in the GLBT domain, it doesn't introduce discernible blockishness to the signal. The GLBT's may be applied in image compression applications to substitute for the forward and inverse DCT. The quantization operation may be applied to the GLBT coefficients in various schemes. Experimental results show improved image quality with less blocking artifact even at a high compression ratio.

B. Generalized Lapped Biorthogonal Transform Embedded Inverse Discrete Cosine Transform

-12-

In an illustrative embodiment of the disclosed system, the front end of the GLBT can be replaced by the DCT. With the factorization in equation (4), the first stage of the GLBT can be written as shown in equation
5   (7).

When the matrices U_{00} and V_{00} are chosen appropriately, the front end of the first stage becomes the DCT. Then the first stage can be written as shown in equation (8), where E_{dct} is a matrix each row of which
10   consists of the DCT basis function. This approach may be taken in order to exploit fast implementations of the DCT available both in software and hardware.

The GLBT with the DCT front end includes the GenLOT, the LOT, and the DCT as special cases for a general
15   choice of K, $K = 2$, and $K = 1$, respectively, with the more strict condition of orthogonality enforced on the matrices U_i's and V_i's.

One advantage of using the DCT front end is to exploit fast and efficient implementations. Another
20   advantage comes in making use of the GLBT in the inversion of standard DCT coefficients. Denoting $G(z)$ as shown in equation (9), the analysis FB and the synthesis FB of the GLBT are shown in equations (10) and (11), where $\overline{G}(z)$ is an appropriate choice for PR. Now
25   considering the FB's pair shown in equations (12) and (13), the analysis FB is the same as the DCT. But the synthesis FB is carried out by completing what's left of the analysis FB in equation (10), followed by the synthesis FB in equation (11). This inverse transform is
30   the GLBT embedded inverse discrete cosine transform (ge-

IDCT). The flowgraph 16 of the forward DCT 18 and ge-IDCT 20 pair is shown in Fig. 2.

The ge-IDCT doesn't look very attractive when processing of the signal is neglected, since the same signal is returned albeit via longer operations. However, the ge-IDCT provides an excellent opportunity to process the signal in the embedded GLBT domain, where the basis functions have much better properties.

C. Deblocking

In applications of lapped transforms, signals may be processed in the lapped transform domains. The disclosed ge-IDCT can be embodied such that the processing of the signal is still in the DCT domain. When the signal is already processed in the DCT domain, the disclosed system can re-process the signal in the embedded lapped transform domain to alleviate harm done by the DCT domain processing.

In particular, the disclosed system may be used to address the blockishness introduced by coarse quantization of the DCT coefficients. In image compression, such coarse quantization results in annoying discontinuity between the data blocks, which is called the blocking artifact. Since the blocking artifact is the result of the independent processing of blocks, it is natural to use information on neighboring blocks in the decoding process to eliminate the blocking artifact. Lapped transforms are excellent examples of such attempts. The ge-IDCT makes neighboring block

information available to a decoding process in the same way that a lapped transform does. The disclosed system uses this neighboring block information to reduce the blocking artifacts.

5   As an example of deblocking, the GLBT with $M = 8$, $L = 2M$, $K = 1$ is now considered, together with the DCT front end. The detailed lattice structure of the GLBT is given as shown in equations (14) and (15). The inverse matrices Ψ_1^{-1} and G_0^{-1} involve transposition of orthogonal

10   matrices and inversion of diagonal matrices, which are trivial. The detailed lattice structures and of the analysis FB and synthesis FB are shown in Figs. 3 and 4, respectively. With M and $L = 2M$, the basis functions of the DCT are M pixels long, whereas the basis functions of

15   the GLBT are L pixels long, ($L - M$) of which overlap.

Nonlinear Weighting

20   Fig. 5 shows how the basis functions of the non-overlapping transforms 30 and the overlapping transforms 32 are interlaid into the entire image. With $M = 8$ and $L = 2M$, the non-overlapping blocks 30 in Fig. 5 are eight pixels long, whereas the overlapping blocks 32 are 16

25   pixels long, with eight pixels overlapped.

       The blocking artifact is a step at the boundary of two adjacent DCT blocks. The location of the step corresponds to the center of the GLBT blocks. Note that the center of the overlapping blocks 32 in Fig. 5 aligns

30   with the boundaries of two adjacent non-overlapping

blocks 30.   The step at this location is going to be represented as a linear combination of odd-symmetric GLBT basis functions.   When M is even, there are $M/2$ even-symmetric and $M/2$ odd-symmetric basis functions.

5   The energy of the odd-symmetric GLBT coefficients may be used as a measure of the blocking effect.   The goal is to detect is a small step due to the blocking artifact.   It can be safely assumed that the energy is fairly small.   Any large amount of energy must be due to

10   real structures in the image.   Hence, the blocking artifact is detected by checking the condition shown in equation (16), where F_K is the kth GLBT coefficient and ϵ_k is the threshold of energy.   The fourth and fifth coefficients correspond to first two odd-symmetric basis

15   functions.   Other odd-symmetric basis functions represent filtering with relatively high pass-bands.   They are excluded for this reason.

        The. blocking artifact has been detected by investigating the energy of the first two odd-symmetric

20   coefficients.   The blockishness is due to small but excessive energy in those coefficients.   The blockishness can be mitigated simply by reducing the energy.   Hence, the odd-symmetric coefficients are weighted with the diagonal weighting matrix shown in equations (17) and

25   (18).   The weighting scheme is nonlinear due to the function α. The use of nonlinear weighting provides selective removal of the blocking artifact without affecting the real structure of the image.   Note that it is still possible that the small energy in F_4 and F_5 is

30   not actually due to the blocking artifact.   In this case,

the shape of the GLBT basis functions along with the fact that the energy is small ensures that no discernible degradation is introduced.

5   The DCT and the ge-IDCT with the disclosed nonlinear frequency weighting can be paired as shown in equations (19) and (20). The ge-IDCT used in place of the IDCT can reconstruct the signal with alleviated blockishness.

Parameter Selection
10

The disclosed nonlinear weighting has only one parameter. It is the threshold of energy ϵ used in detecting the blocking artifact. The threshold is determined as the F_4 value when the input image is as
15   shown in equation (21). This F_4 value corresponds to the energy due to a small step at the adjacent block boundary. Then ϵ can be determined so that one can detect and eliminate the step of δ can be detected and eliminated.

20   The selection of threshold ϵ at various step sizes δ is determined off-line, and the results are stored in a look-up table. The disclosed weighting scheme uses the quality of the reconstructed signal as an input to look up the corresponding threshold ϵ from the table. Hence,
25   the parameter is internal and there is no external parameter that a user must supply. Applications that employ DCT usually have parameters that control the bit-rate and hence the quality. The threshold ϵ can be

chosen in terms of those parameters. For example, *quality factor* in JPEG, QP in H.263, and *mquan* in MPEG can be used in parameter selection.

5   Computational Complexity

The GLBT may be implemented in a fast and efficient manner thanks to the lattice structure. The disclosed ge-IDCT inherits the efficiency of the GLBT. The
10   additional computational complexity imposed by replacing the IDCT with the ge-IDCT is fairly small. Furthermore, some operations can be saved because the weighting is only on odd-symmetric coefficients. For example, the complexity is reduced by half using equations (22) and
15   (23). The matrix multiplication operations can be implemented efficiently by the planar rotations through CORDIC. The disclosed weighting works with various GLBT's. Embodiments may employ integer parameters to reduce the complexity further. Other operations such as
20   W and Λ are trivial. And the operation E'_{dct} in equation (20) is just the IDCT. All the fast implementations of the IDCT, in either software or hardware, are still applicable.

25   D. Design of the ge-IDCT

In this section, the design of an illustrative embodiment of the ge-IDCT is disclosed. The first step is to design a GLBT according to the desired properties.
30   Then the next step is to embed the designed GLBT into the

ge-IDCT.   In designing the ge-IDCT, the following
criteria are considered.

Coding Gain

5

The coding of a transform is defined as shown in
equation (24), where σ^2 is the variance of the input
signal, σ_i^2 is the variance of the ith subband, and $\|h_{si}\|^2$
is the l^2 norm of the ith synthesis filter.   The coding
10   gain measures the energy compaction or decorrelation of
signal from the transform.   In compression applications,
high coding gain is needed so that we can represent an
image with a smaller number of coefficients at low bit
rates. In designing the ge-IDCT, high coding gain helps
15   isolate the frequency components responsible for steps at
the center of the basis functions.

Stopband Attenuation

20   The stopband attenuation is defined as shown in
equation (25).   The stopband attenuation is a classical
criteria for FB design. Low stopband attenuation helps
decorrelation of signal and decreases aliasing between
bands.   Low stopband attenuation also means smooth basis
25   functions.   Consider the ith band filter h_i with a low
pass band.   The Fourier transform of the filter $H_i(e^{jw})$
not only tells us the frequency response of the filter,
but also tells us the shape of the filter's impulse
response, i.e. the basis function.   The lesser the energy

in the stopband, the cleaner the frequency components of the basis function. For the subbands with low pass bands, reducing the stopband attenuations means preventing high frequency components. And hence, the basis functions become smoother.

Smooth basis functions are desired in order to prevent degradation of image quality by the modifications of the lapped transform coefficients. In deblocking, we weight some coefficients. When some of the basis functions are de-emphasized by the weighting, other basis functions become relatively prominent. Any oscillatory behavior of the now-prominent basis functions can degrade image quality.

Design of the ge-IDCT

In designing a ge-IDCT, we desire the following properties. First, we want both the analysis FB and the synthesis FB to have high coding gain such that the signal is decorrelated into specific frequency components. This helps isolate the frequency components responsible for steps at the center of the basis functions. Second, a goal is for both the analysis FB and the synthesis FB to have low stopband attenuations. The reason low stopband attenuations of the FB's is desirable is in part to achieve better decorrelation of the signal. But more importantly, it is to ensure that the basis functions are smooth. The GLBT is designed through the optimization of equation (26), where λ's weight relative importance between the coding gains and

the stopband attenuations. The optimization is over the parameters of the matrices in the lattice structure. Since the GLBT is a biorthogonal transform, the basis functions of the analysis FB and the synthesis FB are

5   different. The GLBT can be designed with different properties for different FB's. In particular, we emphasize the smoothness of the synthesis FB basis functions by trade off between the cost functions through λ's. Once a GLBT with desired properties is designed, it

10   is embedded into a ge-IDCT via equation (20).

E. Experiments

The disclosed ge-IDCT with frequency weighting may

15   be applied to the JPEG still image compression standard. For example, a set of images may be coded by Independent JPEG group's codec at various quality factors, and decoded by the standard JPEG decoder and by the disclosed ge-IDCT with nonlinear weighting. Images may be

20   compressed at quality factors less than 50.

For objective measure, peak signal to noise ratio (PSNR) and mean square difference of slopes (MSDS) may be used. PSNR is given in dB by equation (27), where MSE denotes the mean square error between the original image

25   and the reconstructed image. MSDS is a measure of degradation introduced by the blocking artifacts. The lower the MSDS, the less the degradation due to the blocking artifacts. It should be noted that the energy of the first few odd-symmetric GLBT coefficients is also

30   a good measure for severity of blocking artifacts.

To illustrate improvement in subjective image quality, edge detection may be applied to the compressed and the restored images with the Sobel operator at the same threshold. The resulting edge maps of the image can illustrate removal of artifacts, since the artifact of interest is an undesirable discontinuity.

Applications to JPEG

The PSNR improvement of the ge-IDCT with nonlinear weighting is shown in plots 40, 42, 44 and 46 of Fig. 6 for the airplane, Barbara, Lena, and peppers images respectively. The plots in Fig. 6 show equation (28) in dB. The images decoded by the ge-IDCT show consistent improvement over the images decoded by the standard JPEG at all the quality factors.

The MSDS improvement of the disclosed methods is shown in Fig. 7 for the same images. The plots 50, 52, 54, 56 in Fig. 7 show equation (29), where the negative values indicate reduced MSDS and hence reduced blockishness at block boundaries. As can be seen in Fig. 7, the disclosed methods reduce the MSDS. The results are consistent throughout all the test images at various image qualities. For a quality factor above 50, the threshold ϵ is set at zero. Then the results of the disclosed scheme are identical to the results of the standard JPEG decoder.

Comparative Study

In this section, the results of the disclosed method are compared with existing blocking artifact removal algorithms. The algorithms considered are maximum *a posteriori* (MAP) estimation, and deblocking option in
5 H.263 Annex J. In modeling the prior distribution for the MAP estimation, we use the relation between the robust potential function of the Gibbs distribution and the line process. The relation is given by equation (30), where ρ is the potential function, l is the line
10 process, and Φ_e is the edge penalty function. This relation allows prevention of the discontinuities at the block boundaries explicitly by setting the line process l to one. It should be noted that the MAP estimate is an iterative method with demanding computational complexity.
15 The deblocking filter in H.263 Annex J is tuned for the specific quantization scheme used in H.263. For comparison, images are coded by H.263 I Frame coding method. Then they are decoded with the deblocking option in Annex J and by the disclosed method. The ge-IDCT is
20 designed to cope with the I Frame coding method.

All the methods report similar PSNR and MSDS improvements. But there are distinct differences in subjective image quality. Fig. 8 shows a part of the Lena image compressed by JPEG at quality factor 15. The
25 image 60 in Fig. 8 shows severe blocking artifact, which iscon_rmedby false edges in the edge map 62 in Fig. 8. Fig. 9 shows comparison between the MAP estimation and the disclosed method. Both methods remove the blocking artifact effectively. Differences lie in preservation of
30 details and texture. Because the disclosed method

applies nonlinear weighting only on the specific frequency components, it shows superb preservation of details and texture. The differences are shown clearly on Lena's hat. As shown in Fig. 9, image (a) 70 is the result of MAP estimation, edge map (b) 72 the result of MAP estimation with line process, image (c) 74 the result of the disclosed method, and edge map (d) 76 the result of the disclosed method. Note that most of the texture in Lena's hat is missing in the MAP estimate due to over-smoothing.

Fig. 10 shows a part of the Lena image compressed by the H.263 I Frame coding method at QP = 13. The image 78 and the edge map 80 in Fig. 10 show severe blocking artifacts.

Fig. 11 shows comparison between the deblocking option and the disclosed method. The implementation of the deblocking filter modifies only four pixels near the block boundaries, two pixels on each side. Fig. 11 includes the image (a) 90 by the deblocking filter, edge map (b) 92 generated by the deblocking filter, image (c) 94 generated by the disclosed method, and edge map (d) 96 generated by the proposed method, thus showing that the image processed by the deblocking filter is still relatively blockish due to under-smoothing.

In a smooth region such as Lena's shoulder, the deblocking filter fails to eliminate blockishness. In contrast, the disclosed method modifies the coefficients of the lapped transform basis functions, which are twice the DCT block length, 16 pixels long to be specific. The

disclosed method removes blockishness in smooth regions effectively.

In conclusion, the disclosed system includes the ge-IDCT, that can be paired with the forward DCT. The ge-IDCT inverse transforms the DCT coefficients available at decoders. This aspect is important, because it means there is no incompatibility introduced by replacement of the IDCT by the ge-IDCT. The disclosed inverse transform exploits the lapped transform domain weighing to reconstruct the signal with alleviated blockishness.

The ge-IDCT is based on the lattice structures, which leads to fast and efficient implementation. The additional computational complexity imposed by the new inverse transform is trivial.

Experiments with the JPEG still image compression standard have confirmed the validity of the disclosed transforms. The ge-IDCT has proved to provide better performance than those of complex algorithms at low computational complexity. The ge-IDCT is a competitive alternative to the IDCT in mid to low bit rate still image/video sequence compression applications.

II. Low Bit Rate Video Sequence Coding Artifact Removal

Another illustrative embodiment of the disclosed system is now described, and further describes methods for reducing blocking and ringing artifacts. Consistent with the above discussion, the first technique replaces the conventional inverse DCT (IDCT) of a decoder in order to reduce blockishness. It is referred to herein as the

lapped orthogonal transform embedded IDCT (le-IDCT). The second disclosed technique is a non-linear data adaptive robust filter based on the Maximum Likelihood (ML) model parameter estimation, and is referred to herein as the *robust filter*. The disclosed robust filter is applied to alleviate the ringing artifact.

Advantageously, these two disclosed techniques generally do not require changes in the encoder, or in the bit-stream, and hence may conveniently be standard compliant. Computational complexities of the disclosed techniques are moderate and amenable to real-time implementation within a desktop PC environment.

The disclosed le-IDCT and robust filter are designed carefully such that their use does not degrade major structures of the image. This advantageous property is considered the *robustness* provided. The le-IDCT achieves such robustness by use of selective smoothing through non-linear weighting on only a couple of coefficients. The robust filter achieves its robustness by clustering samples into three clusters and using only the samples in one cluster. Having such robust components as those disclosed herein is beneficial in artifact removal algorithms because it simplifies the way they tab into the decoder. Some of the existing post-processing algorithms use linear filtering to eliminate artifacts. Linear filters may degrade images when they are applied in wrong places. Such existing algorithms have to detect and retain precise locations of artifacts. These additional detecting and book keeping steps significantly

complicate   the   implementation   of   such   existing
algorithms.

Sections II(A) and II(B) below present the le-IDCT
and the robust filter for removal of blocking artifacts
and ringing artifacts, respectively.   In Section II(C)
below, both the le-IDCT and the robust filter are applied
to H.263+ video sequence.   In Section II(D), the
disclosed method is compared to deblocking option of
H.263+ Annex J in terms of picture quality objectively
and subjectively as well as run time complexity.

A. Blocking Artifact Removal

The   blocking   artifact   is   a   consequence   of
independent   processing   of   adjacent   blocks   of   image
pixels.   Better   quality   images   can   be   achieved   by
processing adjacent blocks simultaneously.   Good examples
of simultaneous adjacent block processing techniques are
lapped transforms, in which adjacent processing blocks
overlap each other.   These overlapping transform blocks,
along with the use of gracefully decaying longer basis
functions   ensure   the   reconstructed   image   is   blocking
artifact free even at very low bit rates.
The generalized lapped orthogonal transform (GenLOT)
is   the   general   form   of   lapped   orthogonal   transforms
(LOT's).   The le-IDCT in the disclosed system is based on
the GenLOT.   Essentially, the disclosed system utilizes
the   fact   that   the   first   stage   of   the   GenLOT   can   be
replaced   by   the   DCT   matrix.   Below,   the   GenLOT   is
reviewed, and the le-IDCT described.

The Generalized Lapped Orthogonal Transform

The GenLOT is defined as a linear phase paraunitary
filter bank (LPPUFB) with a polyphase transform matrix
(PTM) given by equation (31). The first stage E_0 is a
LPPRFB with no delay element, and can be factored as
shown in equation (32), where I is the identity matrix
and J is the reversal matrix. The PTM of each stage $G_i(z)$
is given by equation (33).

The matrix $\Lambda(z)$ contains the delay element z^{-1}. The
filter lengths of the GenLOT increase with the delay
element at each ith stage. The matrices U_i and V_i are
orthogonal matrices. The matrices in the PTM in the form
of equivalent Givens rotation angles need to be designed
carefully, and often optimized for better coding gain and
stopband attenuation.

For $K > 1$, adjacent data blocks of the GenLOT
overlap each other. Moreover, the basis functions of the
GenLOT have shapes that decay smoothly to zero. As such,
when an image is processed in the lapped transform
domain, it doesn't introduce discernible blockishness to
the signal. GenLOT's may be applied in image compression
applications to substitute for the DCT. The quantization
operation is applied to the lapped transform coefficients
in various schemes. The results show improved image
quality with less blocking artifacts even at very low bit
rates.

-28-

Lapped Orthogonal Transform Embedded Inverse Discrete
Cosine Transform

5

With appropriate choice of U_0 and V_0 of the GenLOT,
the first stage E_0 becomes the DCT matrix. An apparent
advantage of having the DCT first stage is to exploit
fast and efficient implementation. Another advantage is
to make use of the GenLOT in the inversion of standard
DCT coefficients.

The analysis filter bank (FB) and the synthesis FB
of the GenLOT are shown in equations (34) and (35)
respectively, where $G(z)$ is shown in equation (36).

Consider the FB's pair shown in equations (37) and
(38). The analysis FB is the same as the DCT. But the
synthesis FB is carried out by completing what's left of
the analysis FB in equation (34), followed by a diagonal
weighting matrix A and the synthesis FB in equation (35).
The operation in equation (38) is called the lapped
orthogonal transform embedded IDCT (le-IDCT). Note that
if $A = I$, $R_{le\text{-}IDCT}(z)$ reduces to the usual inverse DCT
matrix.

The le-IDCT provides an excellent opportunity to
process a signal in the embedded lapped transform domain,
where the basis functions have much better properties.

Nonlinear Weighting

The disclosed le-IDCT can be used to eliminate blocking artifacts introduced by coarse quantization of the DCT coefficients. This can be accomplished by choosing appropriate weighting in the diagonal matrix A.

As an example of deblocking, let us consider the GenLOT with $M = 8$, $K = 1$, and the DCT front end. The detailed lattice structure of the GenLOT is given in equations (39) and (40).

Fig. 12 shows an example of the impulse responses 100 and the frequency responses 102 of the GenLOT. This lapped transform can be embedded into the le-IDCT via equation (38).

Let F_k be the kth GenLOT coefficient and ϵ_k be a threshold of energy. The weighing matrix can be chosen as shown in equations (41) and (42). The weighting scheme is nonlinear due to the function α. Use of nonlinear weighting provides selective removal of the blocking artifact without affecting the real structure of the image.

Computational Complexity

The GenLOT has fast and efficient implementation thanks to the lattice structure. The disclosed le-IDCT inherits the efficiency of the GenLOT. Additional computational complexity imposed by replacing the IDCT with the le-IDCT is fairly small. The detailed lattice structure of the le-IDCT is shown in Fig. 13. Some operations can be saved because the weighting is only on odd-symmetric coefficients. The complexity is reduced by

half using equation (43), where A_{odd} is a diagonal matrix
with the weights for only odd-symmetric coefficients.
The matrix multiplication operation can be implemented
efficiently by the planar rotations through CORDIC.

Other operations such as W and Λ are trivial. And
the operation E'_{dct} shown in equation (38) is just the
IDCT.   All the fast implementations of the IDCT, in
either software or hardware, are still applicable. It is
noted that the operation of E'_{dct} is not additional.   It
is an operation a decoder has to perform during the
standard decoding process.   Only the operations that
precede the E'_{dct} are additional.

Parameter Selection

The disclosed nonlinear weighting has only one
parameter.   It is a relatively simple deblocking
algorithm not only in terms of the computations but also
in terms of the number of parameters.

The parameter is the threshold of energy ϵ used in
detecting the blocking artifact.   The threshold is
determined as the F_4 value when the input is as shown in
equation (44).   It is the energy corresponding to a small
step at the adjacent block boundary.   Then ϵ can be
determined such that one can detect and eliminate the
step of δ.

The selection of threshold ϵ at various step size δ
is determined off-line, and the results are stored in a

table. The disclosed weighting scheme uses the quality of the reconstructed signal as an input to look up the corresponding threshold ϵ from the table. Hence, the parameter is internal and there is no external parameter to be supplied by the user. Video compression applications that employ DCT usually have parameters that control the bit-rate and hence the quality. The threshold ϵ can be chosen in terms of those parameters. For example, QP in H.263+ and *mquan* in MPEG can be used in parameter selection.

B. Ringing Artifact Removal

A robust filter is now described to remove mosquito noise as a post-processing approach. Its formulation as an ML estimator and its properties are discussed as follows.

Maximum Likelihood Parameter Estimation

The disclosed system operates by replacing a rippled surface with a flat surface to remove the ringing artifact. The disclosed system attempts to fit a flat surface model to the compressed image as necessary. A flat surface model consists of the number of surfaces, grayscale values of each surface, and corresponding surface information. These parameters are estimated from a given compressed image. In order to manage a broad class of images, a flat surface model is applied locally to small regions of the image.

A $[w \times w]$ window centered at (i,j)th pixel slides through the compressed image g pixel by pixel to pick samples G. Our flat surface model consists of the number of surfaces K, the grayscale values of surface θ, and the surface information z. The surface information z is a $[w \times w]$ matrix with its elements taking the values in $\{1 \ . \ . \ . \ K\}$. The grayscale values of each surface form a $[K \times 1]$ vector θ. The flat surface model image of size $[w \times w]$ can be written as shown in equation (45), where $\mathbf{1}$ is a vector valued indicator function. The center pixel of $\boldsymbol{F}$, denoted by $\boldsymbol{Fc}$, is taken as the (i,j)th pixel of the ringing artifact free image $\hat{f}$. To estimate $\boldsymbol{F}$, we need to estimate z and θ.

The parameter estimation problem is shown in equation (46) where $\boldsymbol{G}$ is incomplete data with z missing. The estimation problem with the complete date $(\boldsymbol{G}, z)$ can be written as shown in equation (47), which can be solved by the k-means algorithm.

A Robust Filter

The number of surfaces K has to be determined from the samples $\boldsymbol{G}$ before the estimation of the probability density $P[\boldsymbol{G}|\theta]$. It can be determined by a hierarchical clustering algorithm with a criterion of merit. A simple alternative is to fix the number of surfaces. A three-cluster model whose cluster centers are determined by a simple rule is used in one embodiment. Given the samples $\boldsymbol{G}$, the cluster centers are initialized as shown in

equation (48), where G_c denotes the grayscale value of the center pixel in the window. Furthermore, the number of iterations in the k-means algorithm is set to one. The estimate is still an ML estimate under the probability density $P[G|\theta]$ approximated by the simplified k-means algorithm.

Note that the center pixel of the window F_c is taken as the (i,j)th pixel of the ringing artifact free image $\hat{f}$. Therefore, θ which the center pixel of F takes is the only parameter of interest. Furthermore, with the simplified k-means algorithm, the result of the three-cluster model is non-iterative in nature. We denote the robust filter as the mapping from the samples G to the estimate F_c. It is robust in the sense that major edge is preserved. This is because pixels belonging to the other side of the edge will be clustered into another cluster and will not be used to estimate the current pixel value.

Let $C(i,j)$ denote the index set of pixels in G centered at (i,j)th pixel. We define the index set $A(i, j; \alpha)$ such that equation (49) holds. For the entire image, the operation of the robust filter is equivalent to equation (50), where $\mathbf{1}$ is the indicator function, $\Sigma_{i,j}$ is over the entire image. Define the function V_c by equation (51). We regard f as a parameter of the function V_c, and adopt a notation $V_c(g; f)$ instead of $V_c(g, f)$. Then the robust filter is the ML estimation of the image with the probability $P[g|f]$ modeled by equation (52). The ML estimate of the parameter f is shown in equation (53), where $\overline{f}^{\alpha}$ is the conditional mean

defined by equation (54), where $\#_{A(i,j;\alpha)}$ is the number of pixels in the set $A(i,j; \alpha)$.

Computational Complexity

Many artifact removal algorithms based on estimation methods are iterative in nature. The computational complexity of iterative algorithms is relatively high. In addition, an image-size buffer is required for intermediate results. As a result, they are not suitable for applications with low complexity as well as low power consumption constraints. The disclosed robust filter is a non-iterative algorithm with low computational complexity. Note that the estimate depends only on the samples in the clique $C(i,j)$. Therefore, it requires only partial information of image and hence a small size buffer.

Another advantage of the disclosed robust filter is its robustness. It removes the ringing artifact without degrading the major structures in image. Consequently it does not need any pre-steps to detect the region with ringing artifact or a carry to convey information through out the decoder. The robust filter can be applied strictly as a post-processing.

C. Experiments

The disclosed techniques have been applied to the coding artifact removal of H.263+ compressed sequences. In one embodiment, the application of the disclosed

system into the decoder is quite simple. The IDCT for I Frames are replaced by the le-IDCT, and the robust filter is applied to every frame as post-processing. The modification of this embodiment is depicted in Fig. 14, wherein "le" 108 represents a part of the le-IDCT that precedes the IDCT, "RF" represents the robust filter, and "s" 112 is a switch.

The test bench is based on H.263+ v3.0 released by The University of British Columbia. A set of test sequences consists of *container*, *foreman*, *hall*, and *news* in qcif format ([144 x 176] frame size). The sequences are compressed at target bit rate of 24 kb/s with I Frames every 100 frames at QP = 13 and QP = 31. The exception is the *foreman* sequence which is compressed at 48 kb/s.

Results of the experiments along with comparison to H.263+, will be discussed in the following section.

Objective Picture Quality

In Fig. 15, the frame-by-frame improvement in PSRN of our disclosed methods over the baseline H.263+ decoder is shown. These results 120 and 122 are for the *foreman* and *hall* sequences at QP = 13 respectively. The plots 120 and 122 show equation (55) for each frame in dB. For the *foreman* sequence, the improvement is moderate. However for the hall sequence, the improvement is consistent for every frame. In Fig. 16, the frame-by-frame improvement in mean square difference of slopes (MSDS) for *foreman* and *hall* sequences at QP = 13 is shown

in plots 130 and 132 respectively. MSDS is a measure of blockishness that gauges the severity of the blocking artifact. The plots 130 and 132 show equation (56) for each frame. Hence, negative values mean improvement over

5   the baseline H.263+. MSDS improvement of *foreman* sequence is on most of the frames, and that of *hall* sequence is on every frame. It reflects reduced blockishness in the sequences processed by the disclosed methods.

10   PSNR and MSDS of the baseline H.263+ and the disclosed method for the test sequences at various QP values are given in Table I 160 of Fig. 19 and Table II 162 of Fig. 20, respectively.

15   Subjective Picture Quality

        Fig. 17 shows a frame 68 of hall sequence at QP = 13. The image (a) 140 in Fig. 17 suffers from both the blocking artifact and mosquito noise. The blocking

20   artifact is most severe in smooth areas of floors and walls, and the ringing artifact is prominent around edges and around the moving person in the center of the frame. The image 142 in Fig. 17 shows effective removal of both artifacts. The edge maps (c) 144 for image (a) 140 and

25   (d) 146 for image (b) 142 in Fig. 17 validate the removal of blockishness in the image.

        D. Comparative Study

In this section, we compare the results of the disclosed system to that of deblocking filter in H.263+ Annex J. For a simple description of the filter, let us consider removal of blocking artifact in a one

5   dimensional case at the first and the second transform blocks. The operation of the filter is written as shown in equation (57), where d1, d2, and clip(.) are designed for appropriate smoothing at different QP.

The deblocking filter of Annex J operates with some

10 other advanced options. These options are not available in a baseline implementation. Both the encoder and the decoder have to be equipped with such advanced options. For fair comparison, the disclosed method is applied with the same options that the deblocking filter uses.

15

Objective Picture Quality

Table III 164 of Fig. 21 and Table IV 166 of Fig. 22 show comparison of methods in PSNR and MSDS. All of the

20 reported numbers are comparable.

Subjective Picture Quality

The shortcoming of the nonlinear filter is that the

25 modifications of pixel values are on four pixels around the block boundaries, two pixels in each side of the boundaries. The effect is to replace the step edge with the graded edge, which result in under-smoothing in smooth regions. The deblocking filter does not provide

30 effective removal of blocking artifact in some sequences.

Fig. 18 shows a frame 68 of hall sequence at QP = 13 with the deblocking option specified in Annex J. The result of the H.263+ deblocking filter is shown in image (a) 150 of Fig. 18. It still shows residue of blockishness even after the deblocking filtering. The image (b) 152 in Fig. 18 is the result of the disclosed method. The same options in Annex J are used except its deblocking filter. The result shows effective removal of the coding artifacts. The edge maps (c) 154 and (d) 156 corresponding to images (a) 150 and (b) 152 respectively in Fig. 18 validate the claim.

Run Time Complexity Comparison

The computational complexity of the disclosed algorithms in terms of run time is investigated. The algorithms are written in straight forward C and embedded into the decoder. They are tested on a 333 MHz dual Pentium PC with 512 MB RAM and SCSI hard-drive running on Windows 2000. The purpose of this comparison is to demonstrate that these algorithms can be applied to H.263+ in real time without assembly coding and human optimization effort. Only speed optimization of Microsoft Visual C++ 5.0 is opted.

The average run time of I and P frames for each sequence are summarized in Table V 168 of Fig. 23 and Table VI 170 of Fig. 24 respectively. The current implementation can decode both I Frames and P Frames at the rate of 20 Frames per second. The frame rates can be

improved further by reducing the overhead due to data movements in the current implementation.

Comparison to the deblocking filter in H.263+ Annex J is also presented in Table V 168 of Fig. 23 and Table VI 170 of Fig. 24. The deblocking filter reports slightly faster run time than the disclosed approaches. It is important to point out that deblocking filter in Annex J is applied not only on the decoder, but also on the encoder which is already overloaded with motion estimation. On the contrary, the disclosed method is designed to work only on the decoder side. It works with the standard bit streams, and higher picture quality is traded off with a little more processing power.

Two coding artifact removal algorithms developed in previous chapters are applied in low bit-rate video sequences. First, the le-IDCT substitutes the IDCT in the decoder. And secondly, the robust filter is applied as post-processing on every frame.

The video sequence coding artifacts of blocking artifact and mosquito nose is suppressed significantly by incorporation of disclosed methods into the decoder.

Additional complexity imposed on either software or hardware implementation of both algorithms is trivial compared to others. They are suitable for communication/storage with constraint bit-rate budget such as H.263+, MPEG2 and MPEG4. Experimental results on motion video show impressive improvement in PSNR, MSDS, and subjective picture quality. In comparative study, the disclosed method is proved to be more effective than the method specified in Annex of the standard.

Those skilled in the art should readily appreciate that programs defining the functions of the disclosed system and method can be implemented in software and delivered to a system for execution in many forms;

5   including, but not limited to: (a) information permanently stored on non-writable storage media (e.g. read only memory devices within a computer such as ROM or CD-ROM disks readable by a computer I/O attachment); (b) information alterably stored on writable storage media

10   (e.g. floppy disks and hard drives); or (c) information conveyed to a computer through communication media for example using baseband signaling or broadband signaling techniques, including carrier wave signaling techniques, such as over computer or telephone networks via a modem.

15   In addition, while the illustrative embodiments may be implemented in computer software, the functions within the illustrative embodiments may alternatively be embodied in part or in whole using hardware components such as Application Specific Integrated Circuits, Field

20   Programmable Gate Arrays, or other hardware, or in some combination of hardware components and software components.

While the invention is described through the above exemplary embodiments, it will be understood by those of

25   ordinary skill in the art that modification to and variation of the illustrated embodiments may be made without departing from the inventive concepts herein disclosed. Accordingly, the invention should not be viewed as limited except by the scope and spirit of the

30   appended claims.

APPENDIX A - Equations

$$E(z) = G_{K-1}(z)G_{K-2}(z)\cdots G_1(z)E_0. \tag{1}$$

$$
\begin{aligned}
E_0 &= \frac{1}{\sqrt{2}}\begin{bmatrix} U_0 & 0 \\ 0 & V_0 \end{bmatrix}\begin{bmatrix} I & J \\ J & -I \end{bmatrix} \\
&\triangleq \frac{1}{\sqrt{2}}\Psi_0 W_0,
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
G_i(z) &= \begin{bmatrix} U_i & 0 \\ 0 & V_i \end{bmatrix}\begin{bmatrix} I & I \\ I & -I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & z^{-1}I \end{bmatrix}\begin{bmatrix} I & I \\ I & -I \end{bmatrix} \\
&\triangleq \Psi_i W \Lambda(z) W.
\end{aligned} \tag{3}
$$

$$\Psi_i = \begin{bmatrix} U_{i1} & 0 \\ 0 & V_{i1} \end{bmatrix}\begin{bmatrix} \Gamma_i & 0 \\ 0 & \Delta_i \end{bmatrix}\begin{bmatrix} U_{i0} & 0 \\ 0 & V_{i0} \end{bmatrix}, \tag{4}$$

$$R(z) = z^{-(K-1)}E_0^{-1}G_1(z^{-1})^{-1}\cdots(z^{-1})G_{K-2}^{-1}(z^{-1})G_{K-1}^{-1}(z^{-1}), \tag{5}$$

such that

$$R(z)E(z) = z^{-(K-1)}I, \quad l \geq 0, \tag{6}$$

$$E_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_{01} & 0 \\ 0 & V_{01} \end{bmatrix} \begin{bmatrix} \Gamma_i & 0 \\ 0 & \Delta_i \end{bmatrix} \begin{bmatrix} U_{00} & 0 \\ 0 & V_{00} \end{bmatrix} W_0. \qquad (7)$$

$$\begin{aligned} E_0 &= \begin{bmatrix} U_{01} & 0 \\ 0 & V_{01} \end{bmatrix} \begin{bmatrix} \Gamma_i & 0 \\ 0 & \Delta_i \end{bmatrix} E_{dct} \\ &\triangleq G_0 E_{dct}, \end{aligned} \qquad (8)$$

$$G(z) = \prod_{i=K-1}^{0} G_i(z). \qquad (9)$$

$$E(z) = G(z) E_{dct} \qquad (10)$$

$$R(z) = E_{dct}^t \widetilde{G}(z), \qquad (11)$$

$$E(z) = E_{dct} \qquad (12)$$

$$R(z) = E_{dct}^t \widetilde{G}(z) G(z). \qquad (13)$$

$$E(z) \;=\; \Psi_1 W \Lambda(z) W G_0 E_0 \tag{14}$$

$$R(z) \;=\; z^{-1} E_0^t G_0^{-1} W \Lambda(z^{-1}) W \Psi_1^{-1}. \tag{15}$$

$$\mathcal{F}_k^2 \le \epsilon_k^2, \quad \text{for } k = 4, 5, \tag{16}$$

$$A(\mathcal{F}; \epsilon) = \text{diag}\left[\alpha(\mathcal{F}_4; \epsilon), \alpha(\mathcal{F}_5; \tfrac{\epsilon}{2}), 1, 1\right], \tag{17}$$

where

$$\alpha(x; \gamma) = \begin{cases} 0.5, & \text{if } |x| \le \gamma \\ 1, & \text{otherwise} \end{cases}. \tag{18}$$

$$E(z) \;=\; E_{dct} \tag{19}$$

$$R(z) \;=\; z^{-1} E_{dct}^t G_0^{-1} W \Lambda(z^{-1}) W \Psi_1^{-1} \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} \Psi_1 \Lambda(z) W G_0. \tag{20}$$

$$[0,0,0,0,0,0,0,0,\delta,\delta,\delta,\delta,\delta,\delta,\delta,\delta]^t. \tag{21}$$

$$\Psi_1^t \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} \Psi_1 = \begin{bmatrix} U_1^t & 0 \\ 0 & V_1^t \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & V_1 \end{bmatrix} \tag{22}$$

$$= \begin{bmatrix} I & 0 \\ 0 & V_1^T A V_1 \end{bmatrix}. \tag{23}$$

$$\Phi_{cg} = \frac{\sigma^2}{\left(\prod_{i=0}^{M-1} \sigma_i^2 \|h_{Si}\|^2\right)^{\frac{1}{M}}}, \tag{24}$$

$$\Phi_{sa} = \sum_{i=0}^{M-1} \int_{w \in stopband} |H_i(e^{jw})|^2 dw. \tag{25}$$

$$\min_{E,R} \{\lambda_{cge} \Phi_{cg}(E) + \lambda_{sae} \Phi_{sa}(E) + \lambda_{cgr} \Phi_{cg}(R) + \lambda_{sar} \Phi_{sa}(R)\}, \tag{26}$$

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right), \tag{27}$$

$$\text{(PSNR of the proposed method)} - \text{(PSNR of JPEG)} \tag{28}$$

$$\text{(MSDS of the proposed method)} - \text{(MSDS of JPEG)}, \tag{29}$$

$$\rho(x) = \min_{l \in [0,1]} \{ x^2 l + \Phi_e(l) \}, \tag{30}$$

$$E(z) = G_{K-1}(z)G_{K-2}(z)\cdots G_1(z)E_0. \qquad (31)$$

$$
\begin{aligned}
E_0 &= \frac{1}{\sqrt{2}}\begin{bmatrix} U_0 & 0 \\ 0 & V_0 \end{bmatrix}\begin{bmatrix} I & J \\ J & -I \end{bmatrix} \\
&\triangleq \frac{1}{\sqrt{2}}\Phi_0 W_0,
\end{aligned}
\qquad (32)
$$

$$
\begin{aligned}
G_i(z) &= \begin{bmatrix} U_i & 0 \\ 0 & V_i \end{bmatrix}\begin{bmatrix} I & I \\ I & -I \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & z^{-1}I \end{bmatrix}\begin{bmatrix} I & I \\ I & -I \end{bmatrix} \\
&\triangleq \Phi_i W \Lambda(z) W.
\end{aligned}
\qquad (33)
$$

$$E_{GenLOT}(z) = G(z)E_{dct} \qquad (34)$$

$$R_{GenLOT}(z) = z^{-(K-1)}E_{dct}^t G^t(z^{-1}), \qquad (35)$$

$$G(z) = \prod_{i=K-1}^{1} G_i(z). \qquad (36)$$

$$E_{le-IDCT}(z) = E_{dct} \qquad (37)$$

$$R_{le-IDCT}(z) = z^{-(K-1)}E_{dct}^t G^t(z^{-1})AG(z). \qquad (38)$$

$$E_{GenLOT}(z) = \Phi_1 W\Lambda(z)W E_{dct} \qquad (39)$$

$$R_{GenLOT}(z) = z^{-1}E_{dct}^t W\Lambda(z^{-1})W\Phi_1^t. \qquad (40)$$

$$A(F;\epsilon) = \text{diag}\left[1,1,1,1,\alpha(F_4;\epsilon),\alpha(F_5;\tfrac{\epsilon}{2}),1,1\right], \qquad (41)$$

where

$$\alpha(x;\gamma) = \begin{cases} 0.5, & \text{if } |x| \le \gamma \\ 1, & \text{otherwise} \end{cases} \qquad (42)$$

$$\Phi_1^t A \Phi_1 = \begin{bmatrix} U_1^t & 0 \\ 0 & V_1^t \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & A_{odd} \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & V_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & V_1^T A_{odd} V_1 \end{bmatrix}, \qquad (43)$$

$$[0, 0, 0, 0, 0, 0, 0, 0, \delta, \delta, \delta, \delta, \delta, \delta, \delta, \delta]^t. \qquad (44)$$

$$\mathcal{F} = \sum_{k=1}^{K} \theta_k \mathbf{1}_{z=k}, \qquad (45)$$

$$\hat{\theta} = \arg\max_{\theta} P[\mathcal{G}|\theta], \qquad (46)$$

$$\hat{\theta} = \arg\max_{\theta} P[\mathcal{G}, z|\theta], \qquad (47)$$

$$\theta = \begin{bmatrix} \mathcal{G}_c - 2\alpha \\ \mathcal{G}_c \\ \mathcal{G}_c + 2\alpha \end{bmatrix}, \qquad (48)$$

$$A(i,j;\alpha) = \{(p,q)|(p,q) \in C(i,j), |g_{p,q} - g_{i,j}| \le \alpha\}.$$

(49)

$$\hat{f} = \arg\min_f \sum_{i,j} \sum_{m,n \in C(i,j)} (g_{m,n} - f_{i,j})^2 \mathbf{1}_{m,n \in A(i,j;\alpha)},$$

(50)

$$V_c(g;f) = \sum_{i,j} \sum_{m,n \in C(i,j)} (g_{m,n} - f_{i,j})^2 \mathbf{1}_{m,n \in A(i,j;\alpha)}.$$

(51)

$$P[g|f] \propto \exp\{-V_c(g;f)\}.$$

(52)

$$\hat{f} = \bar{f}^\alpha,$$

(53)

where $\bar{f}^\alpha$ is the conditional mean defined by

$$\bar{f}_{i,j}^\alpha = \frac{\sum_{p,q \in A(i,j;\alpha)} g_{p,q}}{\#_{A(i,j;\alpha)}}, \tag{54}$$

where $\#_{A(i,j;\alpha)}$ is the number of pixels in the set $A(i,j;\alpha)$.

$$\text{(PSNR of the proposed method)} - \text{(PSNR of baseline H.263+)} \tag{55}$$

$$\text{(MSDS of the proposed method)} - \text{(MSDS of baseline H.263+)} \tag{56}$$

$$\begin{aligned}
\hat{f}_6 &= f_6 - d_2 \\
\hat{f}_7 &= \text{clip}(f_7 + d_1) \\
\hat{f}_8 &= \text{clip}(f_8 - d_1) \\
\hat{f}_9 &= f_9 + d_2,
\end{aligned} \tag{57}$$

CLAIMS

What is claimed is:

5   1. A generalized lapped biorthogonal transform embedded
     inverse discrete cosine transform (ge-IDCT) for providing
     decompression of a compressed still image, comprising:
          receiving a plurality of discrete cosine transform
     (DCT) coefficients from a DCT front end process;
10   applying a non-linear weighting to said compressed
     still image;
          inverse transforming said plurality of discrete
     cosine transform coefficients; and
          reconstructing said still image with alleviated
15   blockishness.

     2. A method for improving the picture quality of video
     frames encoded at relatively low-bit rates by reducing
     the effects of both blocking and ringing artifacts,
20   comprising:
          a first post-processing method including
               applying a lapped orthogonal transform-embedded
          inverse discrete cosine transform (le-IDCT) to allow
          data samples from adjacent blocks to be processed
25   simultaneously, whereby blocking artifacts are
          efficiently mitigated; and
          a second post-processing method including
               applying a non-linear robust filter to reduce
          ringing artifacts.
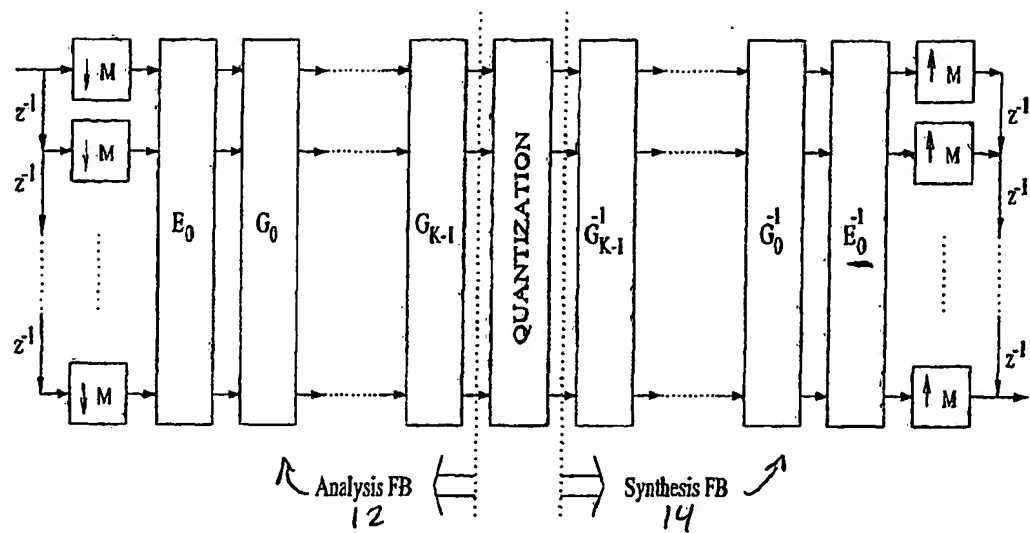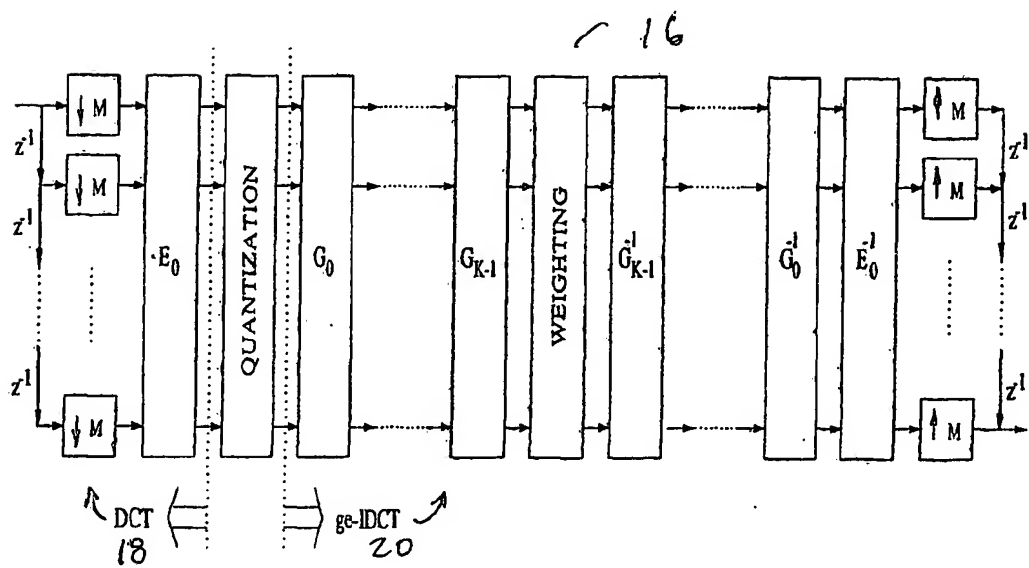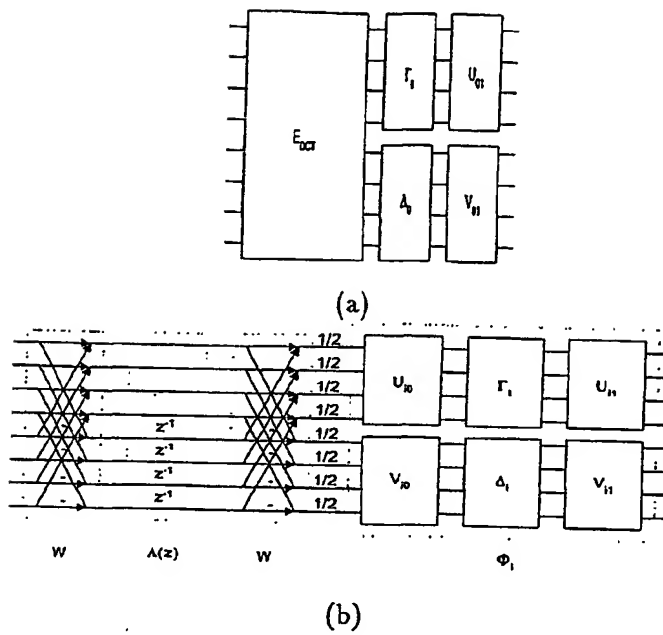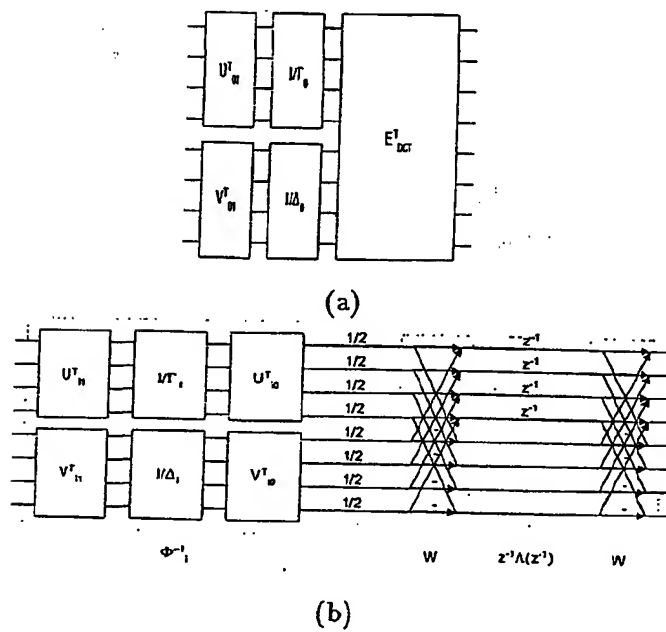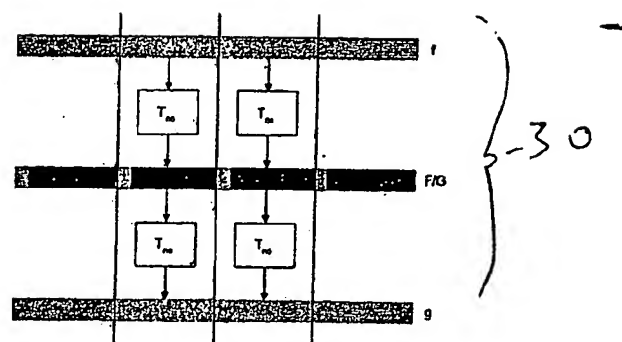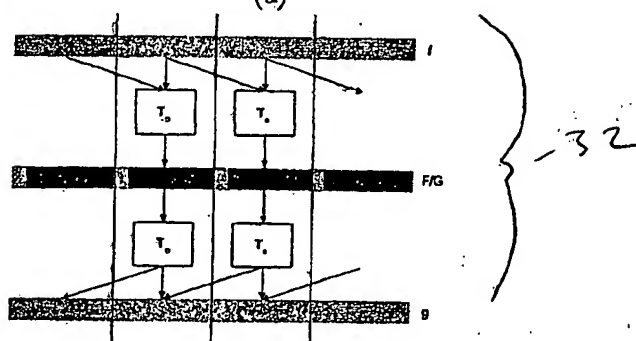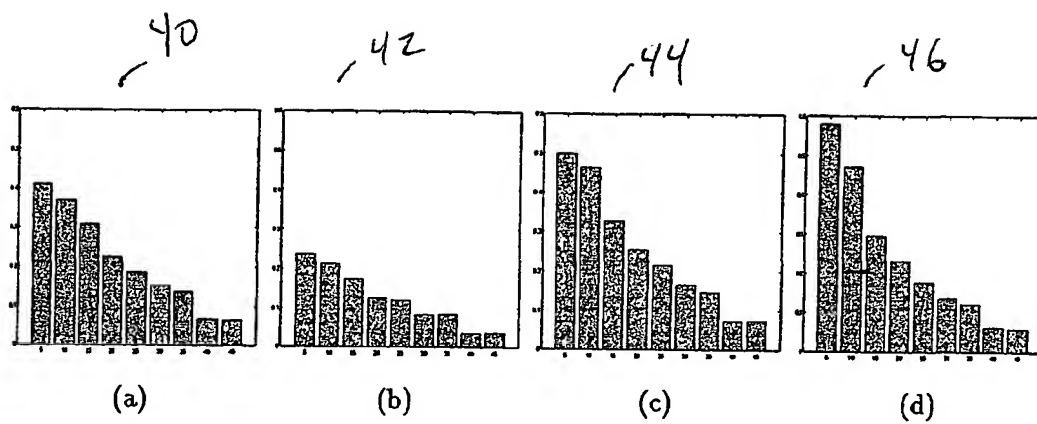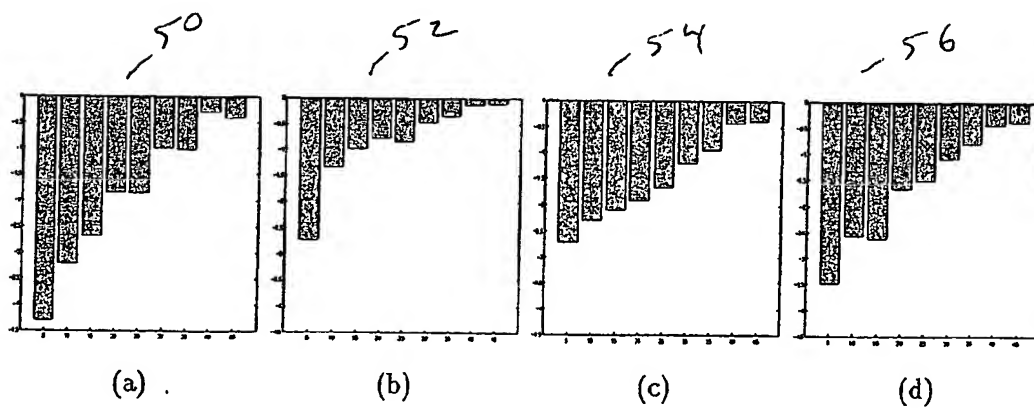
FIGURE 1

FIGURE 2

(a)

(b)

FIGURE 3

(a)

(b)

FIGURE 4

FIGURE 5

FIGURE 6

FIGURE 7

(a)

(b)

FIGURE 8

FIGURE 9

78

80

(a)

(b)

FIGURE 10

(a)

(b)

(c)

(d)

FIGURE 11

(100)   (102)

(a)   (b)

FIGURE 12

FIGURE 13

FIGURE 14

(a)

(b)

FIGURE 15

(a)

~130

(b)

~132

FIGURE 16

140

142

144

146

(a)

(b)

(c)

(d)
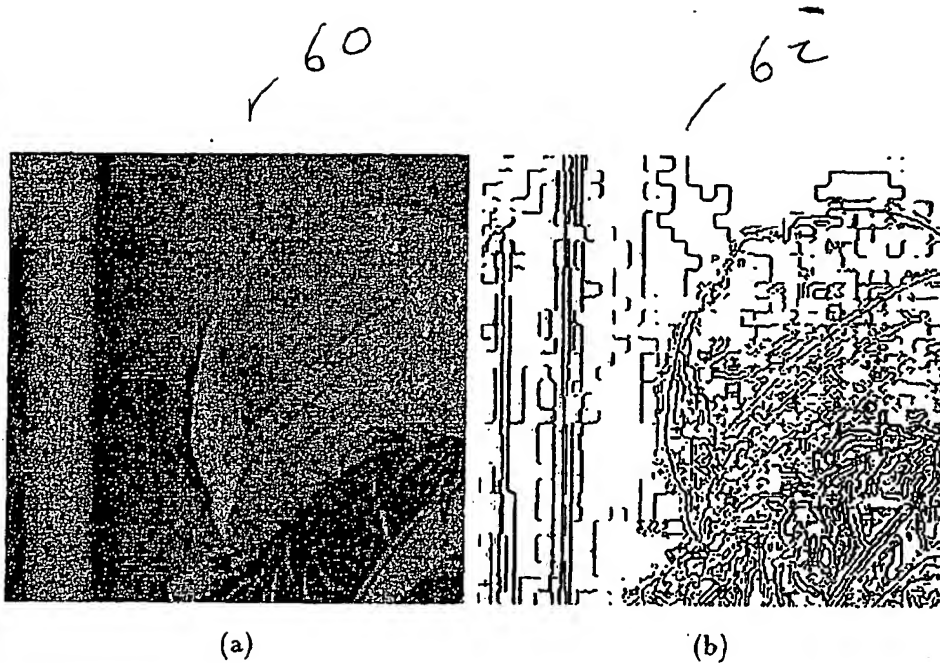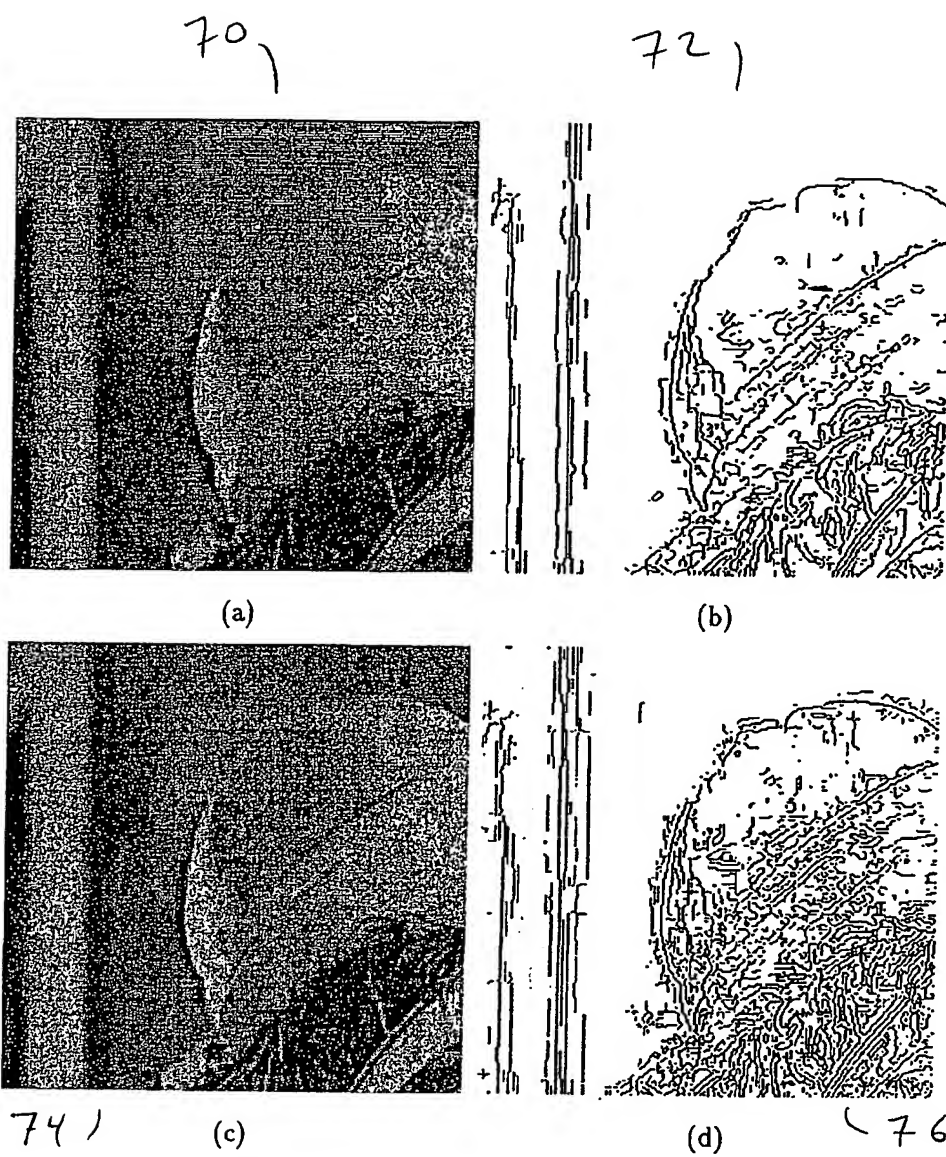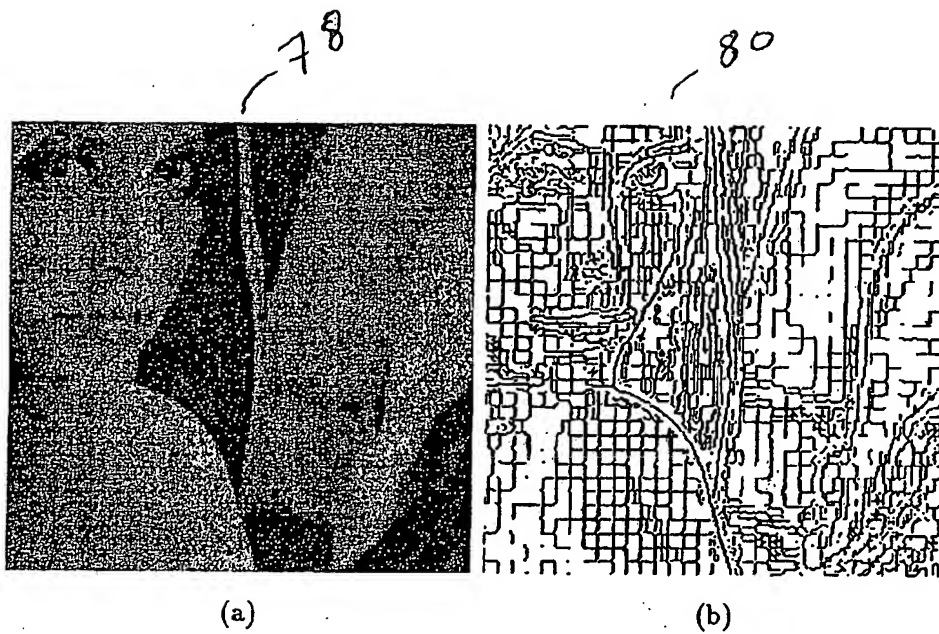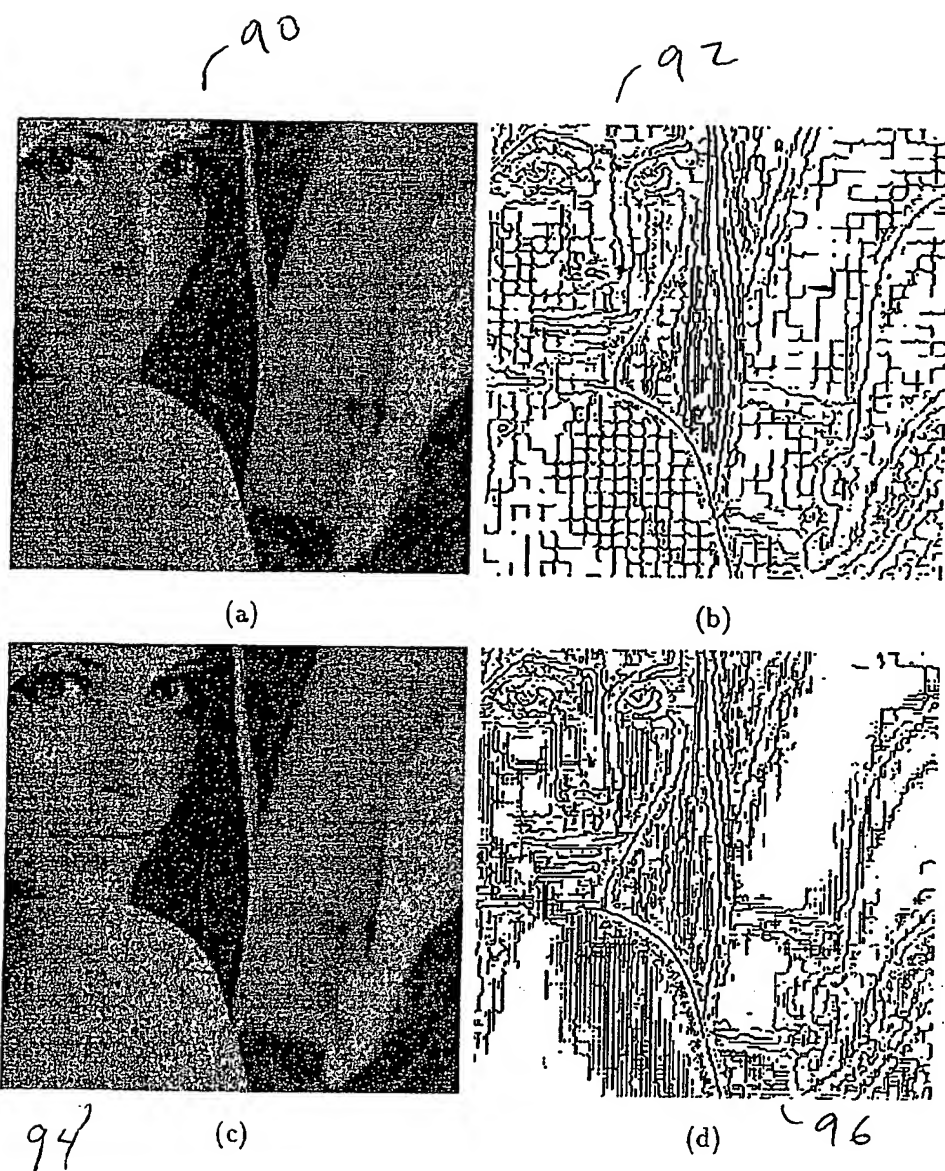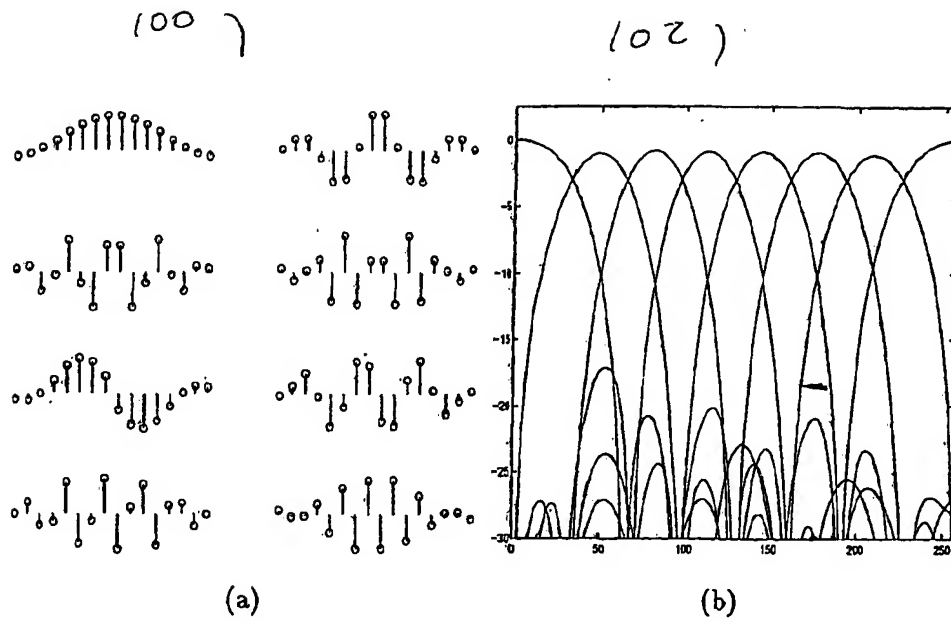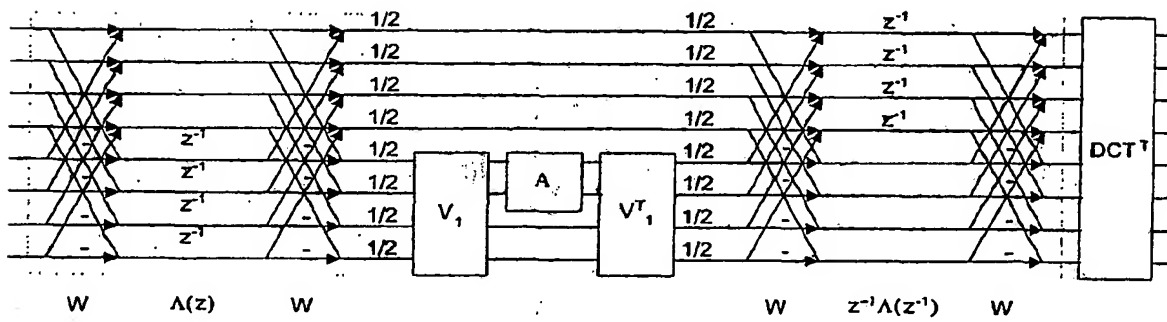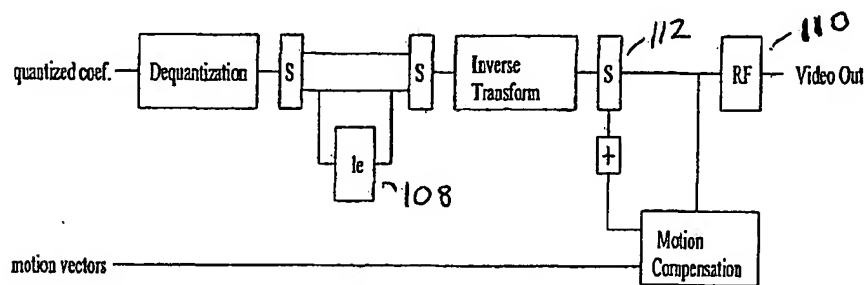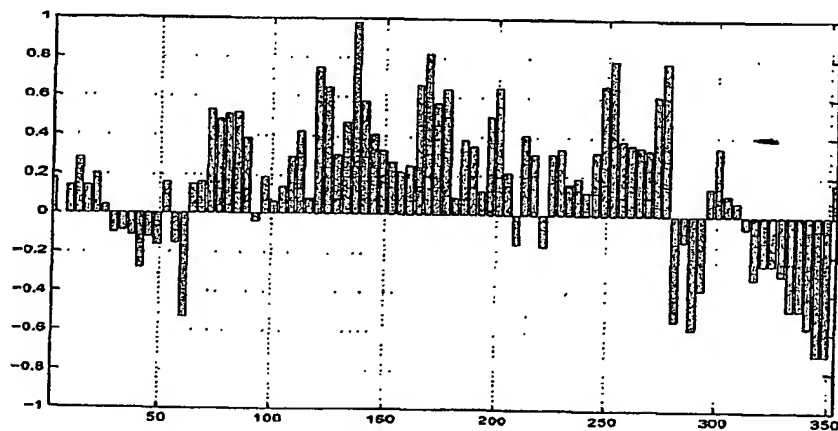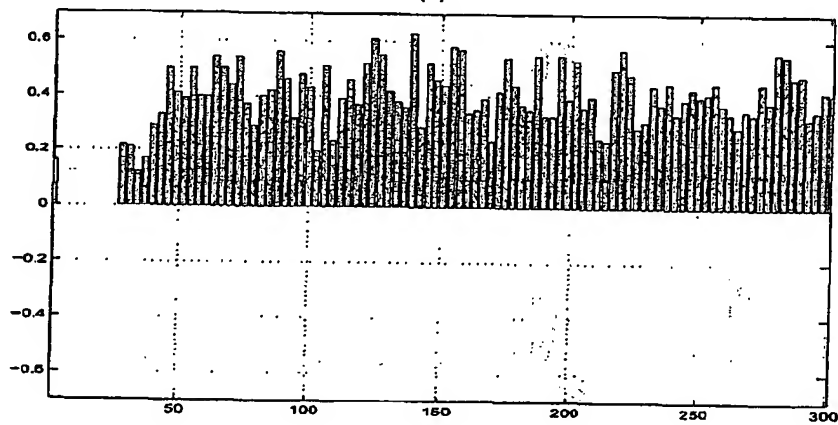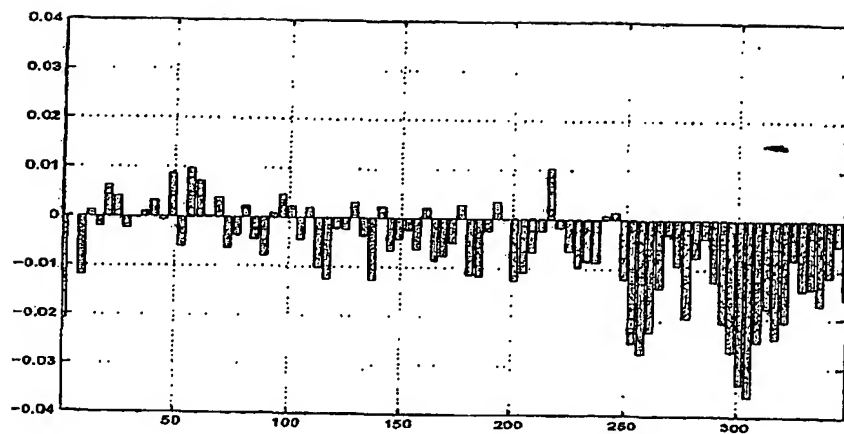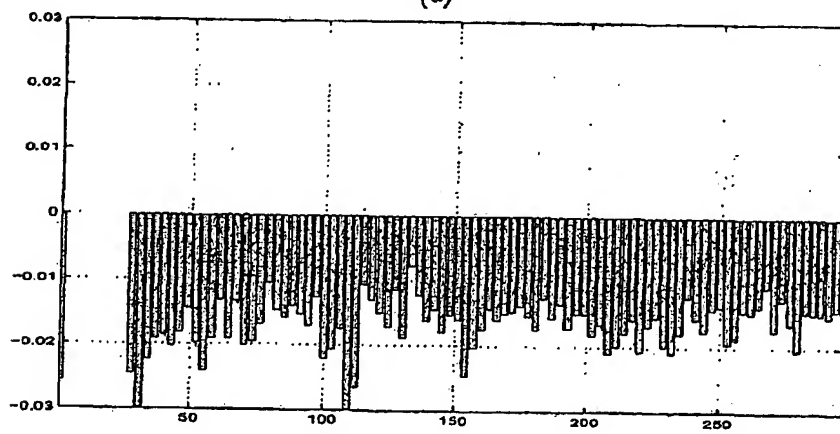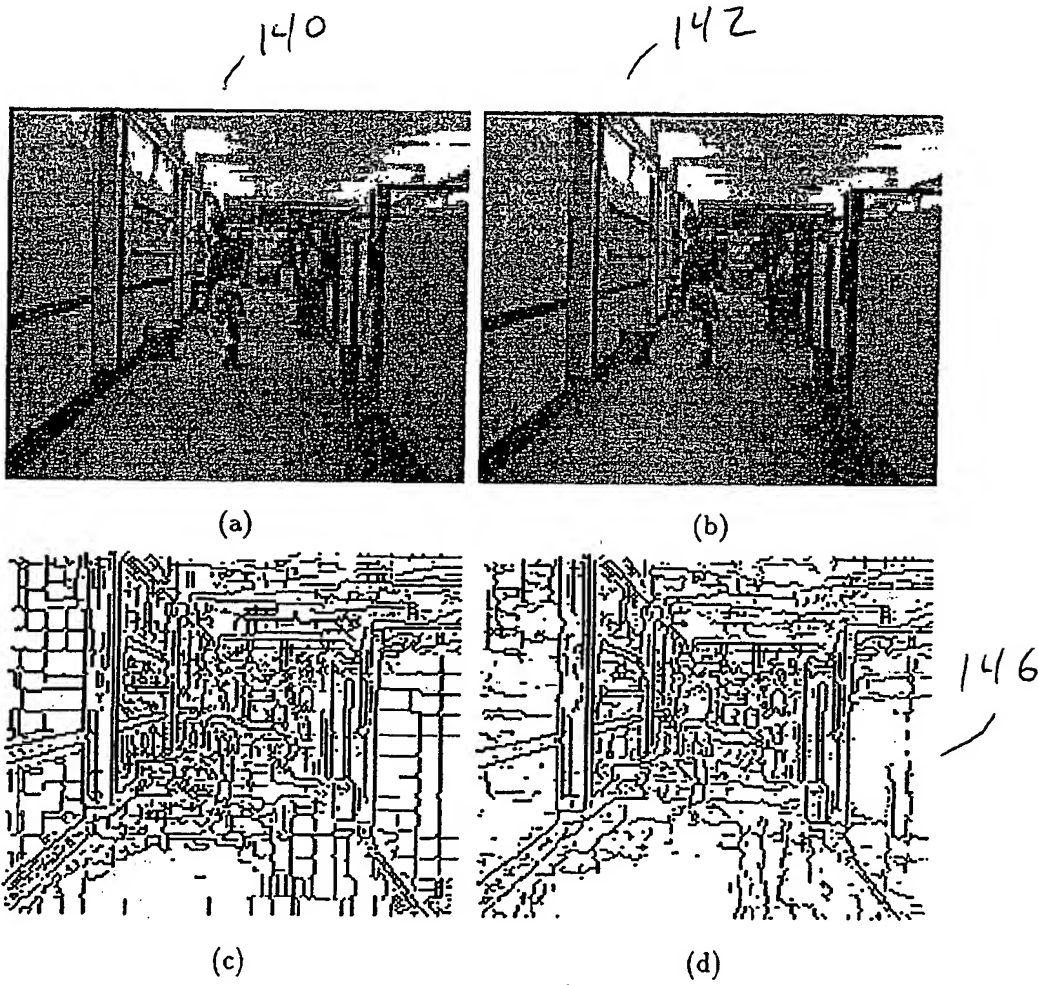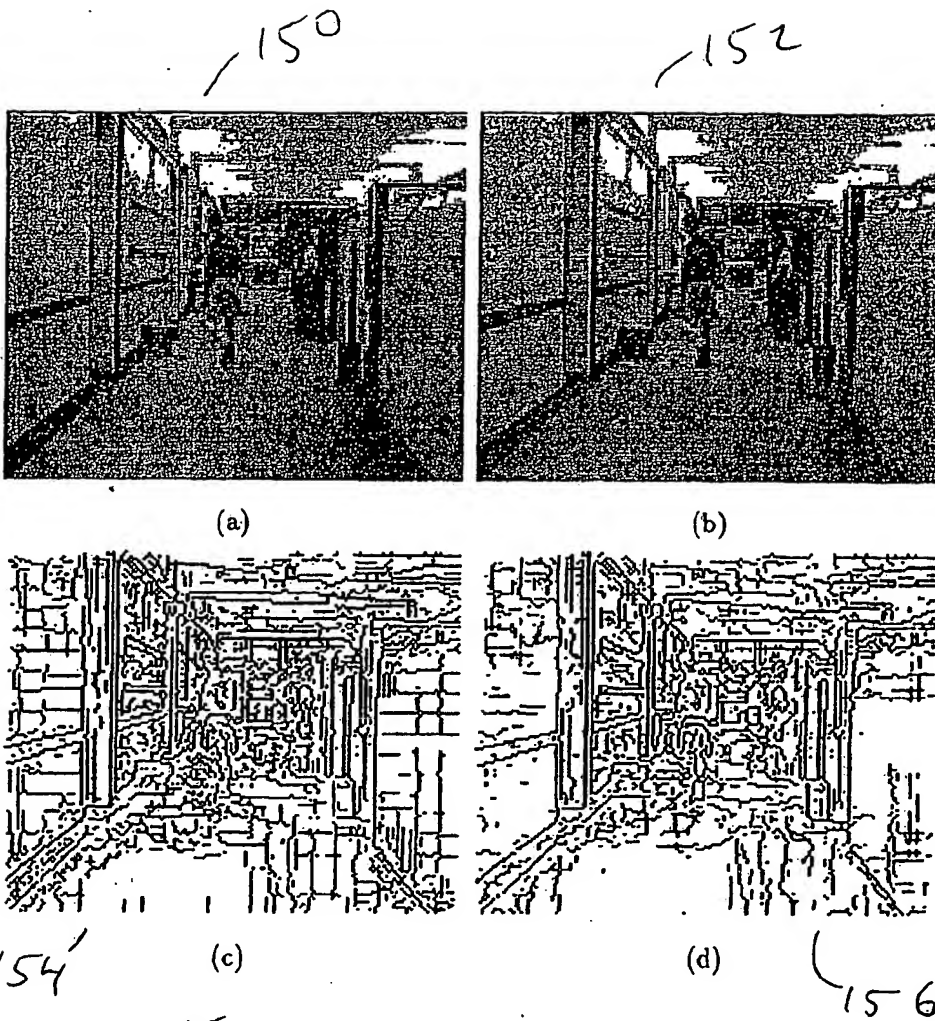
FIGURE 17

FIGURE 18

TABLE I 160

	QP	H.263+	proposed
container	13	32.34	32.19
	31	32.05	31.95
foreman	13	31.63	31.53
	31	31.58	31.50
hall	13	30.18	30.36
	31	29.26	29.45
news	13	29.83	29.89
	31	29.37	29.47

FIGURE 19

TABLE II 162

	QP	H.263+	proposed
container	13	0.36	0.34
	31	0.36	0.34
foreman	13	0.23	0.23
	31	0.24	0.23
hall	13	0.43	0.41
	31	0.43	0.41
news	13	0.41	0.40
	31	0.41	0.39

FIGURE 20

TABLE III 164

	QP	H.263+	DF	proposed
container	13	32.34	32.08	32.20
	31	32.05	31.69	31.93
foreman	13	31.63	32.27	31.80
	31	31.58	32.19	31.74
hall	13	30.18	30.34	30.59
	31	29.26	29.11	29.73
news	13	29.83	29.74	29.90
	31	29.37	29.17	29.50

FIGURE 21

TABLE IV   *(66)*

	QP	H.263+	DF	proposed
container	13	0.36	0.32	0.34
	31	0.36	0.31	0.34
foreman	13	0.23	0.19	0.23
	31	0.24	0.19	0.23
hall	13	0.43	0.39	0.41
	31	0.43	0.36	0.41
news	13	0.41	0.37	0.40
	31	0.41	0.35	0.39

FIGURE 22

TABLE V   *(68)*

	baseline	DF	proposed
container	0.0150	0.0182	0.0489
foreman	0.0143	0.0218	0.0548
hall	0.0142	0.0213	0.0552
news	0.0167	0.0255	0.0583

FIGURE 23

TABLE VI   *(70)*

	baseline	DF	proposed
container	0.0105	0.0165	0.0433
foreman	0.0141	0.0227	0.0454
hall	0.0098	0.0134	0.0417
news	0.0102	0.0154	0.0417

FIGURE 24

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :H04N 7/12

US CL :375/240.22; 382/276, 280, 250

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 375/240.22; 382/276, 280, 250

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,101,279 A (NGUYEN et al.) 08 August 2000, col. 2, lines 1-67.	1-2
A	US 5,850,482 A (MEANY et al.) 15 December 1998, ALL	1-2
A	US 5,838,377 A (GREENE) 17 November 1998, ALL	1-2

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"E"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 OCTOBER 2001	16 NOV 2001

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	GEORGE A. BUGG
Facsimile No. (703) 305-3230	Telephone No. (703) 305-2326

Form PCT/ISA/210 (second sheet) (July 1998)★